AUTOMATED SCREENING OF METABOLIC DISORDERS USING PATTERN RECOGNITION OF GC-MS FULL SCAN SPECTRA FROM URINE ORGANIC ACIDS

by

David K. Crockett

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Medical Informatics

University of Utah

August 2002

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

SUPERVISORY COMMITTEE APPROVAL

of a thesis submitted by

David K. Crockett

This thesis has been read by each member of the following supervisory committee, and by majority vote has been found to be satisfactory.

William L. Roberts

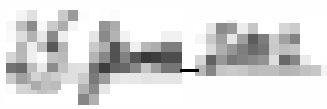
Edward R. Ashwood

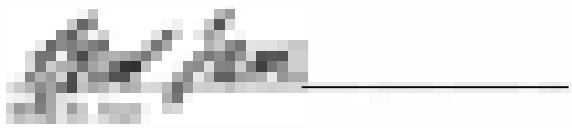
THE UNIVERSITY OF UTAH GRADUATE SCHOOL

FINAL READING APPROVAL

To the Graduate Council of the University of Utah:

I have read the thesis of David K. Crockett in its final form and have found that (1) its format, citations, and bibliographic style are consistent and acceptable; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the supervisory committee and is ready for submission to The Graduate School.





Chair, Supervisory Committee

Approved for the Major Department

Reed M. Gardner
Chair/Dean

Approved for the Graduate Council

David S. Chapman
Dean of The Graduate School

ABSTRACT

Analysis of organic acids in urine is a valuable tool in the diagnosis of the inborn errors of metabolism known as organic acidurias. This test is commonly ordered in newborns with symptoms such as lethargy, failure to thrive, hepatic failure, and suspected familial disorders. A drawback of published methods is the overwhelming amount of data to examine for each patient, prior to the final laboratory report. Physicians will wait as long as two weeks for these time critical results. The goal of this research was to develop an expert system to automate the process of screening for metabolic disorders of urine organic acids.

The Xaminer® pattern recognition software (ThermoFinnigan, San Jose, CA) was adapted to predict and identify patterns of urine organic acid disorders. The gas chromatography-mass spectrometry (GC-MS) full scan spectra of organic acids were used to build the pattern match library and train the software to recognize methylmalonic aciduria (MMA) and associated vitamin B12 deficiency, as well as, a subset of fatty acid oxidation defects (FAOD), including medium chain acyl-CoA dehydrogenase (MCAD) deficiency. Patient data files were de-identified and reprocessed using the expert system. The expert system results were compared to the original laboratory findings.

From a total of 2573 samples, the original laboratory findings were 20 positives for MMA and 29 positives for FAOD. The Xaminer software identified 17 of the 20 MMA positives, plus 4 additional candidate samples that matched the search pattern

criteria. The software found 26 of the 29 FAOD positives. Five additional samples were found to be candidates for FAOD. Software analysis time averaged less than 10 seconds per sample.

This expert system can use pattern recognition of full scan GC-MS data to aid in patient screening for MMA and fatty acid oxidation disorders. The performance of Xaminer shows promise for refining or expanding the reference library to include other metabolic disorders as well.

TABLE OF CONTENTS

ABSTRACT..... iv

LIST OF TABLES..... vii

ACKNOWLEDGEMENTS..... viii

INTRODUCTION..... 1

METHODS..... 4

RESULTS..... 9

DISCUSSION..... 18

CONCLUSIONS..... 21

REFERENCES..... 22

LIST OF TABLES

Table		Page
1.	Xaminer match results for MMA and FAOD samples	10
2.	Xaminer results for MMA and FAOD at varying match thresholds	11
3.	Xaminer results at varying match thresholds for unknown samples	12
4.	Xaminer results at varying training set size for MMA	13
5.	Time requirements for chromatographic interpretation of urine organic acids	17

ACKNOWLEDGEMENTS

This work was supported by the ARUP Institute for Clinical and Experimental Pathology (Salt Lake City, UT). Technical development of the sample preparation and GC-MS assay by Mark Kushnir, Ai Ping Liu, and Leah Richardson, ARUP Laboratories (Salt Lake City, UT). Assistance in chromatographic interpretation and gracious use of the mass spectral libraries, CHLA and HSC, provided by Dr. Dennis Lehotay, Saskatchewan Department of Health (Saskatchewan, Canada).

INTRODUCTION

Metabolic diseases arise from inherited errors of biological pathways. The symptoms are normally due to a buildup of a biosynthetic precursor or product that the body cannot excrete. Although individual metabolic diseases are rare, their cumulative effect is quite common. One recent review states that 1 in 3,500 infants born in the United States is affected by an inborn error of metabolism (1).

Pathologic changes in normal catabolism of amino acids, carbohydrates, biogenic amines, and steroids often results in abnormal excretion patterns of organic acids. These disorders manifest themselves either by excretion of organic acids that are normally not present in healthy individuals, or excretion of massive amounts of acids which are normally present only in low concentrations. Common indications for organic acid screening include clinical situations such as an acute life-threatening episode, metabolic acidosis or hyperammonemia, failure to thrive, recurrent vomiting, neurological deterioration, or following therapy for a specific inborn metabolic error.

Laboratory methods for identifying such disorders utilize gas chromatography (GC), high performance liquid chromatography (HPLC), gas chromatography-mass spectrometry (GC-MS), and liquid chromatography-mass spectrometry (LC-MS) instruments. The principal method for urinary organic acid analysis remains GC-MS. Both liquid-liquid and solid-phase extraction techniques have been used for sample cleanup. The majority of methods employ derivatization to convert the organic acids into

their volatile derivatives. Instrument run times for previously reported GC-MS methods range up to 60 minutes or longer per sample (2,3).

Recent advances in the field of capillary gas chromatography, such as microbore columns, derivatization, fast temperature programming, and improved computer performance allows a significant reduction in analysis time without compromising chromatographic resolution. Current laboratory methods are simple, fast, reproducible, and do not require frequent instrument maintenance. Over 150 clinically important organic acids can be identified in a 13-minute run (4).

The remaining limitation of current methods, however, is the time required to analyze the large amount of clinical data gathered in each sample run. The interpretation of the organic acid screen data is very labor intensive and requires extensive training and experience. Only a few laboratories offer this type of testing, and require a resident expert on staff to report results back to the client. Physicians will often wait as long as two weeks for a final urine organic acid report.

The goal of this project was to develop an expert system to automate the process of identifying metabolic disorders from urine organic acid profiles. The software was intended to aid in interpretation, not to replace human judgement. This is a very complex assay, and involves very demanding interpretation.

The scope of this project was to adapt and evaluate an existing pattern recognition software package for utility in matching organic acid disorders. The Xaminer® software (ThermoFinnigan, San Jose, CA) was originally designed to perform in the setting of arson investigation. It was developed to identify the GC-MS patterns of accelerants and

suspected compounds used to start a fire. The aim of this study was to test the feasibility of using the same software algorithms to detect patterns of urine organic acids by their GC-MS spectra, and to diagnose metabolic disorders by using this pattern recognition. In order to narrow the focus of the project, a representative single acid disorder - methylmalonic aciduria (MMA), and multiple acid disorder - fatty acid oxidation defects (FAOD) were evaluated. If the software was shown to perform well on this initial set of disorders, the Xaminer pattern library could then be expanded to cover a more comprehensive set of organic acidurias. This expert system would aid in the labor intensive process of chromatographic interpretation of the organic acid disorders in urine.

Because the intent of the pattern match software was to serve as a screening tool, the set of positives used to build the Xaminer pattern library contained samples with concentrations of methylmalonic acid equal to 80 μmol/L and higher. This concentration of methylmalonic acid is more consistent with vitamin B12 deficiency and not a true methylmalonic aciduria (5). Pattern library entries for the fatty acid oxidation disorders, included patterns for varying types of fatty acid oxidation disorders, including short chain (SCAD), MCAD, and long chain (LCHAD) acyl-CoA dehydrogenase deficiencies.

METHODS

Samples submitted for the testing were analyzed for creatinine and the volume to be extracted adjusted according to sample creatinine concentration. This was done by calculating aliquot volume = 11.3 / concentration of creatinine in mg/dL. Thus, the extraction volume for each specimen was adjusted to contain 1 μmole of creatinine. Because the injection volume of each sample was the same used for the calibration, the concentration of an acid in each sample can be quantified in mmol/mol creatinine.

Internal standard and hydroxylamine were added to the calibrator, control, and patient samples. Water was added to the samples to bring the volume to 10 mL. The samples were then transferred into anion exchange solid phase extraction (SPE) columns preconditioned with methanol and water. The SPE columns were washed to remove coextracted extraneous materials, dried, and the extracted compounds were eluted. The eluent was evaporated, and the residues derivatized. To eliminate the formation of multiple peaks and to prevent the possible decarboxylation of alpha ketoacids during extraction, a two-step derivatization of oxime and silyl ester formation was used. After derivatization, each sample was transferred to an autosampler vial, and the aliquots were injected into the GC-MS. A representative GC-MS full scan chromatogram of urine organic acids is shown in Figure 1.

The analysis was performed using a Finnigan Voyager GC-MS, equipped with an A220S autosampler and Xcalibur™ software (Thermoquest, San Jose, CA).

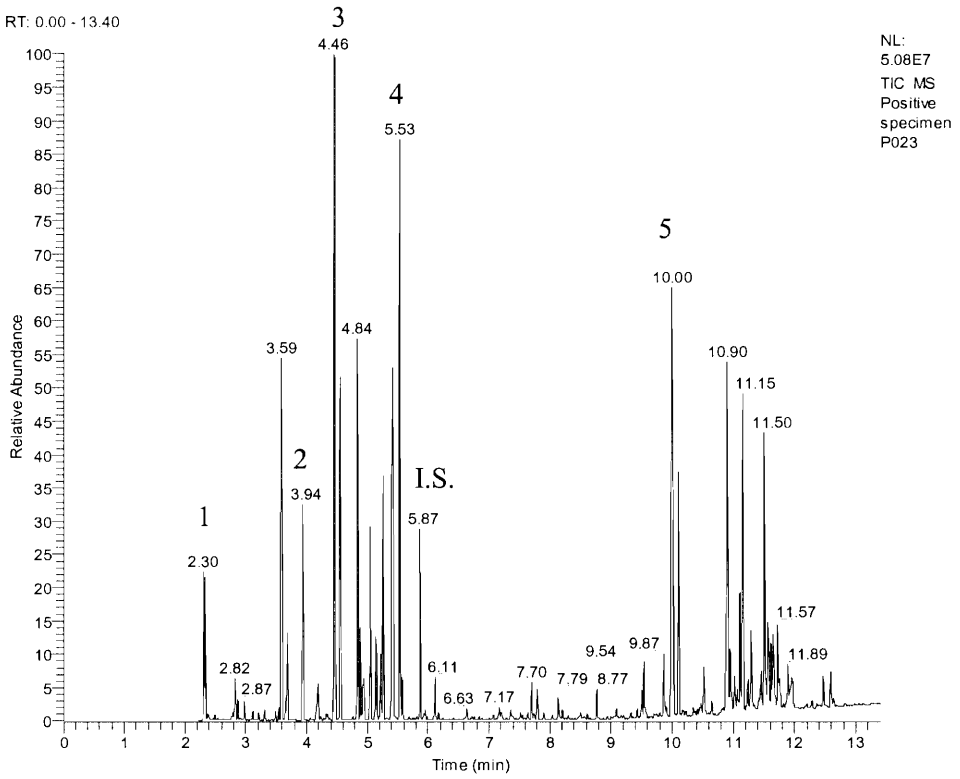


Figure 1. Gas chromatography-mass spectrometry (GC-MS) full scan chromatogram of extracted patient sample showing organic acids in urine. The labeled peaks are: (1) lactic acid, (2) methylmalonic acid, (3) levulinic acid, (4) acetylglycine, and (5) arachidonic acid. The internal standard (2-oxohexanoic acid) is at retention time 5.87.

Data acquisition was in both full scan and selected ion monitoring (SIM) mode using electron ionization with the detector at 350 mV. The GC was equipped with an RTX-1 capillary column, 20 meter length, 0.18 mm diameter, and 0.40 μm film thickness (Restek, Bellafonte, PA). The instrument was tuned using perflurotributylamine.

Data analysis was performed with two methods: full scan and SIM. The full scan method selected the 80 largest peaks in a chromatogram, integrated the peaks, performed a library search, and identified the three best matches in the libraries for each of the peaks in the unknown mass spectra. In addition to the full scan data analysis, quantitation for a total of 36 organic acids was also performed using different SIM programs. Prior to reporting test results, the data for each sample are evaluated first, by a staff technician, and then by the laboratory medical director. Human interpretation was assumed to be the gold standard.

Since Xaminer relies on accurate peak identification and match quality of full scan mass spectral data, the GC-MS full scan library was a crucial step in the software process. Two existing libraries of organic compounds were available for use. A third library was custom built by full scan analysis of organic acid reference standards. The three GC-MS full scan libraries combined contain mass spectral data for more than 400 unique compounds

Data files (n = 2573) for which urine organic acid screening had been performed by GC-MS, were collected for 18 months. All patient data files were de-identified before reprocessing by the Xaminer software. This research was approved by the Institutional Review Board of the University of Utah Health Sciences Center (IRB #9102).

After GC-MS data acquisition, Xaminer performed its own analysis to "fingerprint" or match and report the identity of complex samples. Software training was done by using the data files of previously assayed patient samples with known disorders. Sets of positive samples were chosen for both MMA (n = 20) and FAOD (n = 29) and the chromatographic data entered into the Xaminer software pattern library. Pattern recognition parameters of peak weighting and match quality threshold were optimized for both MMA and the fatty acid oxidation disorders. The expert system was then tested against the entire set of 2573 data files. The goal of the project was to readily identify a predefined set of disorders, including a single acid disorder (MMA) and a multiple acid disorder (FAOD).

The results from the expert system were compared to the previously reported lab findings. The accuracy of identifying unknown organic acids and the precision (reproducibility) of the expert system were studied. Pattern match thresholds of 95%, 85%, 65%, and 50% were observed for both MMA and FAOD. In a separate experiment, a new set of unknown samples (n = 400) was processed against the Xaminer pattern library and the results recorded. Using the set of 400 unknowns, the size of the training set (n = 2, 5, and 20) was also studied for MMA. An evaluation of time saved by Xaminer interpreting the GC-MS full scan data was also produced. Laboratory statistics commonly used for evaluation of assay performance are displayed in Figure 2.

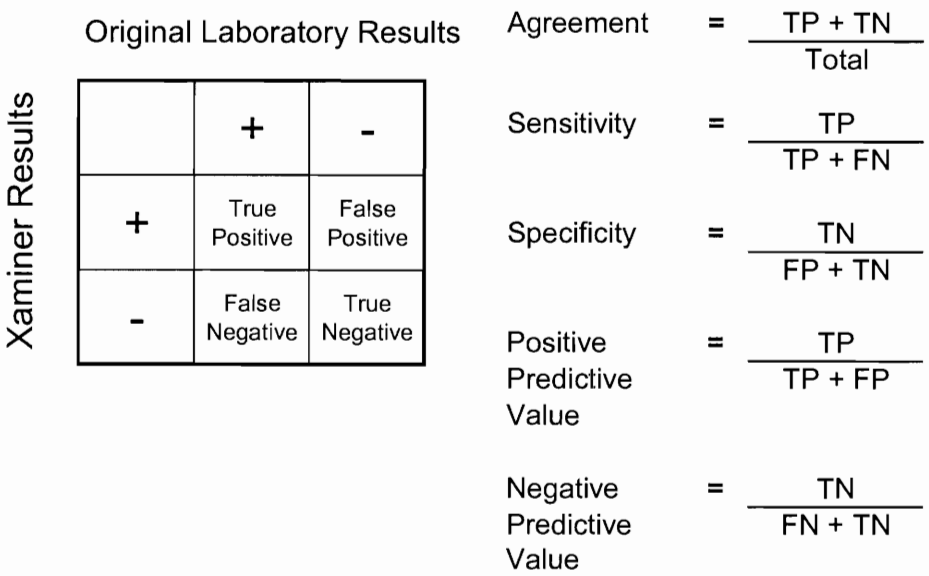


Figure 2. Laboratory statistics used for evaluation of assay performance.

RESULTS

In the original 2573 samples, laboratory findings were 20 positives for MMA and 29 positives for fatty acid oxidation disorders. Using a pattern match threshold of 80%, the Xaminer software identified 17 of the 20 MMA positives, plus 4 additional candidate samples that matched the search pattern criteria. Table 1 shows the detailed performance of the Xaminer software resulting in 17 true MMA positives identified, with 4 false positives and 3 false negatives. Again, using a pattern match threshold of 80%, the Xaminer software found 26 of the 29 FAOD positives in the set of 2573 samples. Five additional samples were found to be candidates for fatty acid oxidation disorders. Table 1 also shows the summary of the software resulting in 26 true FAOD positives identified, with 5 false positives and 3 false negatives. To insure the accuracy of the pattern match results by Xaminer, the discrepant false positives and false negatives were independently verified.

Results for both MMA and fatty acid oxidation disorders with varying thresholds of pattern match quality (95%, 80%, 65%, and 50%) are displayed in Table 2. Results for the set of 400 unknown samples are summarized in Table 3, with pattern match thresholds of 80%, 65%, 50%, and 10% compared. The results of the MMA training size set (n = 2, 5, and 20) is shown in Table 4. Xaminer pattern matching for MMA is shown in Figure 3, and for MCAD in Figure 4.

Table 1. Xaminer match results for MMA and FAOD samples.[1]

MMA	Laboratory[2]	Xaminer[3]
True Positives	20	17
True Negatives	2553	2549
False Positives		4
False Negatives		3

FAOD	Laboratory[2]	Xaminer[3]
True Positives	29	26
True Negatives	2544	2539
False Positives		5
False Negatives		3

[1] Total samples was 2573.
[2] The laboratory findings are assumed to be the gold standard.
[3] Pattern match threshold of 80% was used.

Table 2. Xaminer results for MMA and FAOD at varying match thresholds.[1]

MMA[2]	95% match	80% match	65% match	50% match
Agreement	99.4	99.6	99.0	98.4
Sensitivity	38.1	80.9	80.9	100.0
Specificity	100.0	99.8	99.1	98.3
+ Prediction	88.8	80.9	44.7	32.7
- Prediction	95.7	99.8	99.8	100.0
FAOD[3]	**95% match**	**80% match**	**65% match**	**50% match**
Agreement	99.2	99.6	97.1	95.9
Sensitivity	41.3	89.6	93.1	100.0
Specificity	99.9	99.8	97.1	95.9
+ Prediction	85.7	83.8	27.2	21.9
- Prediction	99.3	99.8	99.9	100.0

[1] Total samples was 2573.
[2] Xaminer results for MMA as compared to the original laboratory findings.
[3] Xaminer results for FAOD as compared to the original laboratory findings.

Table 3. Xaminer results at varying match thresholds for unknown samples.[1]

MMA[2]	80% match	65% match	50% match	10% match
Agreement	98.7	97.0	96.7	79.2
Sensitivity	66.7	77.7	88.9	100.0
Specificity	99.4	97.4	96.9	78.7
+ Prediction	75.0	41.1	40.0	9.7
- Prediction	99.2	99.7	99.7	100.0
FAOD[3]	80% match	65% match	50% match	10% match
Agreement	98.7	98.2	97.3	71.2
Sensitivity	50.0	50.0	75.0	75.0
Specificity	99.4	98.7	97.4	71.7
+ Prediction	40.0	28.5	23.1	2.6
- Prediction	99.4	99.4	99.5	99.6

[1] Total samples was 400, including 9 MMA and 4 FAOD cases.
[2] Xaminer results for MMA as compared to the original laboratory findings.
[3] Xaminer results for FAOD as compared to the original laboratory findings.

Table 4. Xaminer results at varying training set size for MMA.[1]

MMA[2]	set = 2	set = 5	set = 20
True Positives	3	4	8
True Negatives	389	388	379
False Positives	2	3	12
False Negatives	6	5	1

[1] Total samples was 400 unknowns.
[2] Pattern match threshold of 10% was used.

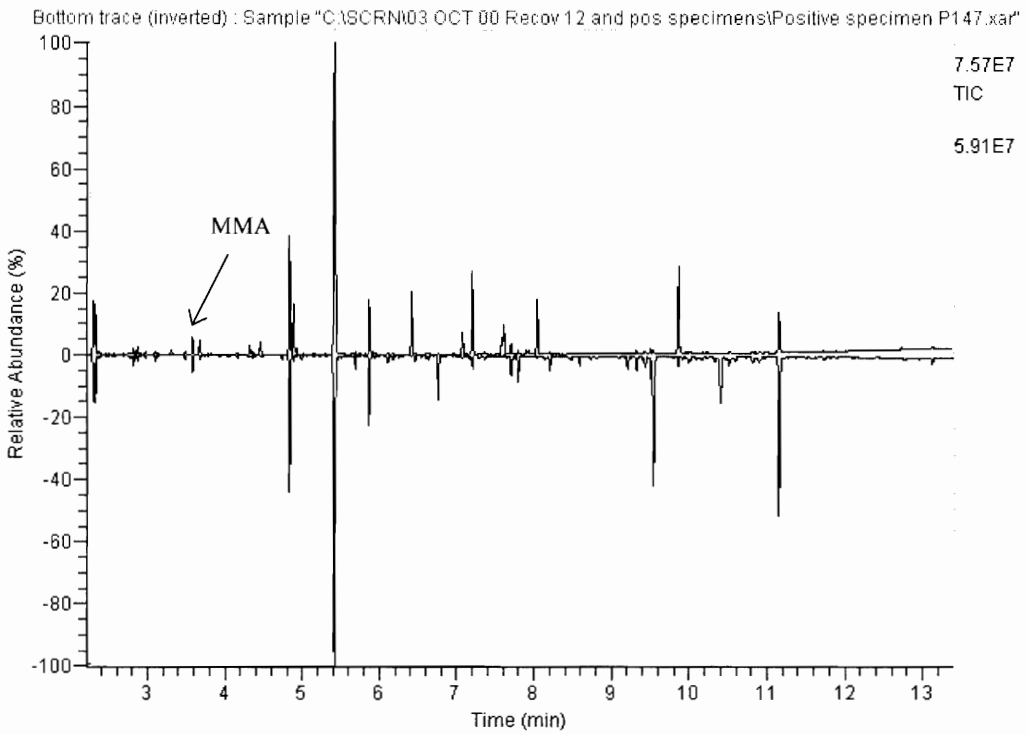


Figure 3. Xaminer spectrogram and inverted pattern match for methylmalonic aciduria (MMA).

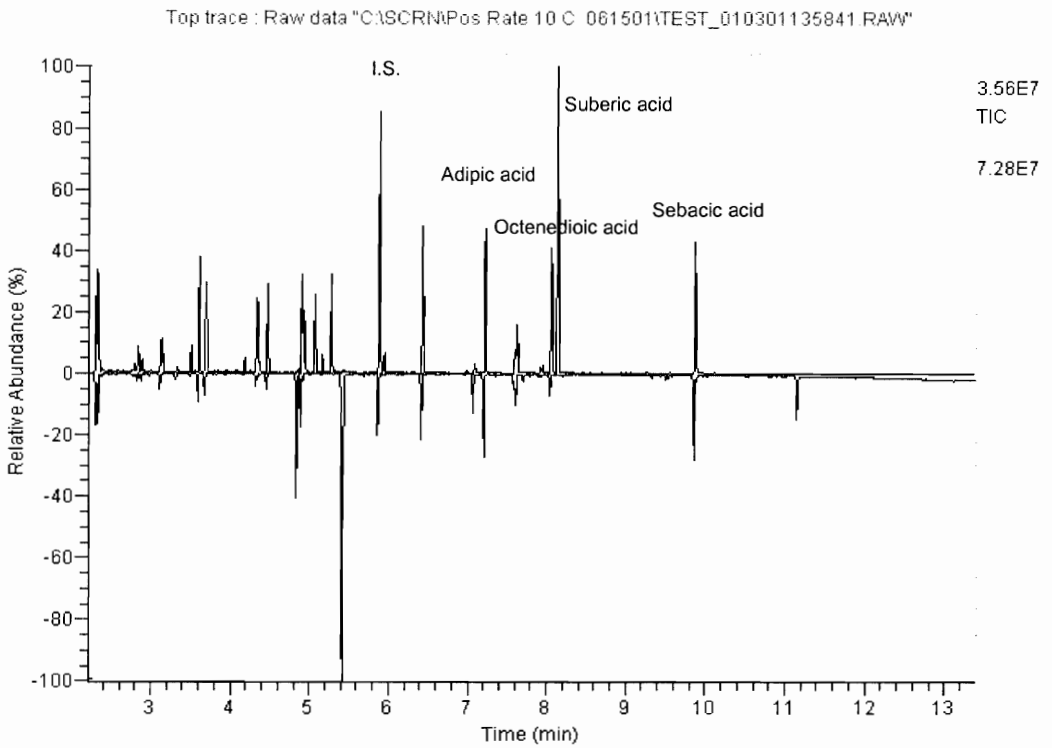


Figure 4. Xaminer spectrogram and inverted pattern match for medium chain acyl-CoA dehydrogenase (MCAD) deficiency.

The average daily workload in the clinical laboratory was 20 to 30 samples. Data analysis for each patient took between 20 to 30 minutes. The medical director spent another 10 to 15 minutes per patient report. The overall time spent per each sample was therefore 30 to 45 minutes. For a run of 20 patient samples, an average of 10 hours of chromatographic interpretation was required prior to releasing laboratory results. The Xaminer software could complete the pre-screen for MMA and fatty acid oxidation disorders in less than 10 seconds per patient sample. These results are summarized in Table 5.

Table 5. Time requirement for chromatographic interpretation of urine organic acids.[1]

Interpretation	Laboratory[2]	Xaminer[2]
Time require for 1 positive sample	45 minutes	10 seconds
Time require for 1 negative sample	15 minutes	10 seconds
Time for 20 samples	10 hours	< 5 minutes

[1] Expert human interpretation is assumed to be the gold standard.
[2] Reported times are averaged for daily workload of approximately 20 samples.

DISCUSSION

The Xaminer® pattern recognition software is a tool to provide an automated data collection, library building and chromatographic pattern matching solution for GC-MS users. Typical applications of the Xaminer software have been arson investigation, perfumes and fragrance characterization, seized drug authentication (cocaine, heroine, etc.), forensic paint analysis, characterization of bacteria and yeasts, and oil spill pollution monitoring. The real world applications are generally associated with large numbers of samples containing many common compounds but in varying abundance. There may be thousands of samples, with hundreds of components, each with very minor differences in relative abundance. This task often requires time intensive expert data review and can be aided by computerized pattern matching of extracted ion chromatograms.

The Xaminer software allows identification of complex mixtures by matching chromatogram traces of unknown samples with known reference standards. Xaminer automates and accelerates pattern matching by allowing the user to compile a library of reference materials. Each material is analyzed for target components and relative intensity information is compiled to construct a fingerprint or "spectrogram" of the sample. An unknown sample may then be searched against this library, based on its own automatically generated spectrogram. Detected component peaks are mapped as arrays of relative abundances, but no longer tied to a time axis. Instead, each peak is allocated a

component number by Xaminer. This forms Xaminer's sample fingerprint for a specific mixture. Due to the proprietary nature of the software, technical details of the actual pattern searching and match algorithms are not readily available.

The overwhelming amount of data per injection remains a significant challenge to routine laboratory analysis. The same pattern algorithm can be mathematically applied to each patient chromatogram. This is more objective than the human eye scrutinizing each sample report, and therefore produces consistent, nonsubjective results. Since 80% to 90% of samples that are submitted for testing are determined to be negative, samples that were screened negative by the expert software, could be directly reviewed by the medical director and results reported on the same working day. It should noted, however, that the comparison of time saved was an estimation of Xaminer searching for two types of disorders only, while the human counterpart was looking for all plausible disorders.

Interesting to note was the effects of lowering Xaminer's pattern match threshold setting. As the match threshold was decreased, a higher percent of false positives was produced. However, to err on the side of caution, a laboratory would likely choose to operate at a lower match threshold to insure that the initial screen identified all positive samples. In order to recognize all positive MMA and FAOD samples in the training sets, Xaminer required a match threshold setting of 50%. Using the 400 unknowns, match threshold settings were decreased in 10% increments until 100% sensitivity for MMA was achieved. However, 100% sensitivity for FAOD was not attained. Training set sizes of $n = 2$ and $n = 5$ did not sufficiently represent positive MMA samples. The training set of $n = 20$ improved the MMA pattern match performance.

It would be impossible to predict that Xaminer would perform well to identify every known organic acid disorder, including children of differing age and nutritional condition. This would require that the Xaminer pattern library contain perhaps multiple entries for every possible combination of variables such as patient age, drug therapy, and nutritional status for each known metabolic disorder (6). However, assuming that such a complete library could be constructed in Xaminer, a remaining complication is the inability of Xaminer to identify new or yet unknown disorders. Such a disorder would likely be misidentified as "normal" and disregarded as clinically insignificant.

One possible strategy is to train the Xaminer system to screen for normal or negative samples, instead of looking for all positive acid disorders. This "rule out" screening strategy has been successful in cervical cytology, and may prove useful in the diagnoses of metabolic disorders as well (7).

CONCLUSIONS

Analysis of organic acids in urine is a valuable tool in the diagnosis of the inborn errors of metabolism known as organic acidurias. The need to separate, detect, and quantify a large number of organic acids results in long instrument run times of 40 to 60 minutes per sample. An improved GC-MS method can identify more than 150 important urine organic acids in 13 minutes. However, the remaining limitation of published procedures is low throughput, due to the labor-intensive identification of organic acids in each sample chromatogram.

The overall performance of the Xaminer software was surprisingly good. However, the inability to reliably screen with 100% sensitivity, precludes this system from being useful in a clinical laboratory setting. Notable, was the large majority of samples (80% to 90% negative) that can be quickly screened and reported by the laboratory. The smaller subset of presumptive positive aciduria samples could then be the focus of the laboratory staff. Unfortunately, without known patterns entered into the library, new and yet unknown disorders would likely be misclassified as normal. This expert system can use pattern recognition of full scan GC-MS data to aid in patient screening for MMA and fatty acid oxidation disorders. The performance of Xaminer shows promise for refining or expanding the reference library to include other metabolic disorders.

REFERENCES

1. Willis RC. Diagnosing newborns. Modern Drug Discovery 2002;5(4):28-33.

2. Tuchman M, Ulstrom RA. Urinary organic acids in health and disease. Adv Pediatr 1985;32:469-506.

3. Niwa T. Metabolic profiling with gas chromatography-mass spectrometry and its application to clinical medicine. J Chromatogr 1986;379:313-345.

4. Kushnir MM, Liu AP, Crockett DK, Urry FM, Roberts WL. A GC-MS method for organic acids screening in urine. Poster #1906 at PittCon 2001.

5. Scriver CR, Beaudet AL, Sly WS, Valle D. The Metabolic and Molecular Bases of Inherited Disease. 8th ed. New York:McGraw-Hill. 2001.

6. Kumps A, Duez P, Mardens Y. Metabolic, nutritional, iatrogenic, and artificial sources of urinary organic acids. Clin Chem 2002;48(5):708-717.

7. Minge L, Fleming M, VanGeem T, Bishop JW. AutoCyte Prep system vs. conventional cervical cytology. Comparison based on 2,156 cases. J Reprod Med 2000;45(3):179-84.